BBC RD 1974/2

1969
THE QUEEN'S AWARD TO INDUSTRY

RESEARCH DEPARTMENT REPORT

An experimental investigation of test-noise signals for the measurement of non-linear distortion of sound signals

R.A. Belcher, B.Sc.(Hons.)

Research Department, Engineering Division
THE BRITISH BROADCASTING CORPORATION January 1974

BBC RD 1974/2
UDC 534.86
621.391.82

AN EXPERIMENTAL INVESTIGATION OF TEST-NOISE SIGNALS FOR THE MEASUREMENT OF NON-LINEAR DISTORTION OF SOUND SIGNALS

R.A. Belcher, B.Sc. (Hons.)

Summary

Ideally it should be possible to estimate by an objective measurement the degree of programme quality impairment caused by circuit non-linearity. The report describes investigations into the correlation between objective measurements of distortion and subjective assessments of programme impairment, using circuits where non-linearity increases with programme signal level. A proposal is given for a method of distortion measurement which uses a test signal formed from pseudo-random noise. This type of measurement, referred to as a noise-separation test, gives better correlation with subjective assessments than the total harmonic distortion test used at present.

Issued under the authority of

P. Rainger

Head of Research Department

Research Department, Engineering Division,
BRITISH BROADCASTING CORPORATION

January 1974

(EL-84)

AN EXPERIMENTAL INVESTIGATION OF TEST-NOISE SIGNALS FOR THE MEASUREMENT OF NON-LINEAR DISTORTION OF SOUND SIGNALS

AN EXPERIMENTAL INVESTIGATION OF TEST-NOISE SIGNALS FOR THE MEASUREMENT OF NON-LINEAR DISTORTION OF SOUND SIGNALS

R.A. Belcher, B.Sc.(Hons.)

1. Introduction

When non-linearity is present in sound signal circuits, the quality of the reproduced programme may be impaired. It would be convenient if some objective measure of the non-linearity could be used to estimate, with reasonable accuracy, the degree of subjective impairment that the programme would suffer.

At present, most sound broadcasting organisations test for this form of non-linearity by measuring the 'total harmonic distortion' (t.h.d.) of the circuit. A sine-wave test signal is used, and the t.h.d. is expressed as the ratio of the r.m.s. level of harmonic components to the r.m.s. level of test signal measured at the output of the circuit. A t.h.d. measurement is sometimes quoted as a %, but in this report it will be quoted as a decibel difference. The t.h.d. test is easy to instrument, and is used mainly for checking the day-to-day performance of circuits which have been assessed subjectively by listening for programme quality impairment. It is widely recognised, however, that t.h.d. measurements do not necessarily give a good indication of the subjective impairment due to non-linearity. Ways of modifying the t.h.d. test have been examined by various authors in the hope that a test more suitable for type-approval purposes might be formulated.

Shorter,[1] and later Wigan,[2] proposed as a test method spectral analysis, and subsequent weighting of the harmonic distortion products of a 1 kHz test signal. The weighting factors suggested by the two authors were different, but both proposed that increased account should be taken of the amplitude of higher-order harmonics. Wigan's proposal for a Weighted Distortion Criterion took into account more accurately the subjective impairment produced by high-order harmonics. However, both these proposals referred *only* to non-linear circuits for which the total harmonic distortion was substantially independent of signal level. In practice, non-linear distortion of the form most common in sound signal circuits generally increases with signal level; these proposals are therefore not considered further in this report.

Non-linearity measurements with a 1 kHz test signal are inadequate for specifying non-linear distortion as no account is taken of the effect of intermodulation products, nor of variations of non-linearity with signal frequency. In an effort to take more account of these parameters, more recent investigations have used a test signal formed from random noise.

In 1970 the CCIR proposed for study a standard test signal:[3] random noise which was weighted to have a power spectrum similar to that of an average programme. It was intended primarily for use in measuring cross-talk between carrier circuits; although suggested as being suitable, also, for measuring non-linear distortion, no method was suggested for the latter application. Since the CCIR noise signal occupies most of the audio spectrum it is not really appropriate for distortion measurement.

One practical method, studied by various broadcasting organisations[4,5] requires the removal of part of the spectrum of the standard test signal. Distortion products which fall in the resulting spectral gap are selected by a suitable filter, and their level registered on a meter. BBC variants of the standard test signal are examined in this report, and they are referred to as test-noise signals. The difference in level between the total unfiltered test-noise signal measured at the output of the circuit under test and the selected distortion products, expressed in dB, is termed the *noise-separation.* The test method will be referred to as the *noise-separation* method to distinguish it from the more familiar measurement of total harmonic distortion.

The present investigation was initiated in the BBC Designs Department, but the bulk of the work, including corroborating the initial findings, was carried out at the Research Department. For preliminary investigations the spectrum of the CCIR standard test signal was restricted to an upper-frequency limit of 3·5 kHz. Informal subjective tests provided data relating the threshold of impairment to programme signal level for three distorting media: two BBC line receiving amplifiers, and a saturating transformer. The results of this preliminary investigation corroborated the Designs Department findings, and helped to identify the most promising line of investigation for future work.

The main investigation included formal subjective tests aimed at providing accurate subjective data relating programme impairment to programme signal level. The programme items used were more critical of non-linear distortion than those items used initially, and the tests employed a wider range of non-linear circuits in order to generate more types of non-linear distortion. *Noise-separation* measurements were made with various test-noise signals to find whether any would give an objective measurement which agreed well with subjective assessments for the types of distortion investigated.

The investigations described in this report relate to circuits in which non-linear distortion increases with applied signal level. A formal subjective investigation provided data relating programme impairment to signal level for three programme items and four non-linear circuits. This data was used to judge the ability of distortion measuring equipment to indicate the subjective impairment to be expected for a range of conditions. In addition it was used to find optimum values of several important parameters for the *noise-separation* test.

2. Preliminary investigations

2.1. General

Section 1 has outlined the requirements for a reliable non-linearity test, and has suggested that a random noise signal may be more suitable than a sine-wave for use as the test signal. A preliminary investigation was conducted to see if a variant of the CCIR standard test signal could give estimates of the threshold of impairment for non-linear distortion, equal to or better than those obtainable from measurements of t.h.d.

Informal subjective tests were used to estimate (a) programme levels at which each of two BBC line receiving amplifiers caused just noticeable impairment of an excerpt of piano music, and (b) the programme level at which a saturating transformer produced just noticeable impairment of an excerpt of organ music.

These preliminary investigations were a basis for the work described in later sections of this report and revealed the main problems associated with providing reliable objective and subjective estimates of non-linear distortion.

2.2. Method of measuring noise-separation

In general the *noise-separation* test for measuring non-linearity requires apparatus for generating a test-noise signal, further apparatus for selecting the products of non-linear distortion, and a meter for indicating the level of distortion products relative to the level of the test-noise signal. Fig. 1 is a simplified block diagram which illustrates how the equipment is interconnected with the circuit under test.

The generator filter (2) in Fig. 1, weights the spectrum of the noise source (1) to form the test-noise signal. This filter (2) comprises two networks in tandem: (a) a network which filters the noise into frequency-bands, and (b) a network which gives frequency-dependent weighting to the noise signal.

The input gain of the circuit under test is adjusted to produce the desired amount of non-linear distortion; the output gain of the circuit is adjusted to give unity gain between points A and B, under linear operating conditions.

Input and output impedances are 600Ω. Programme level at point A would normally be zero programme volume: nominal programme peaks register 8 dB above a line-up of 0 dBm when measured on a Peak-Programme Meter (PPM).[6] The detector filter (3) is a similar arrangement to the generator filter (2): one network selects the distortion products, and the other network provides frequency-dependent weighting.

The meter (4) is first calibrated to read 0 dB when monitoring the signal level at B; *noise-separation* is then measured by monitoring the signal level at the output of the detector filter (3).

2.3. The investigation

Most of the power of the CCIR standard test signal is present in the frequency band below 3·5 kHz as the average programme weighting is −10 dB at this frequency. This frequency was used as an upper limit for the bandwidth of the first test-noise signal as this choice also leaves a fairly wide bandwidth for measuring distortion products. The upper frequency limit of the test-noise signal is defined by the 3·5 kHz low-pass filter response shown in Fig. 2(*a*), and distortion products are selected by the 4·5 kHz high-pass filter response shown in Fig. 2(*b*). No frequency-dependent weighting was given to these products in the preliminary investigations. Two meters were used for measuring *noise-separation*: (a) a PPM and (b) an r.m.s. meter of the thermocouple type.

Noise-separation was measured with various test-noise powers in the range 0 dBm ± 6 dBm measured at A, and each non-linear circuit was operated such that zero programme volume at A would produce just perceptible impairment of the programme as defined by the informal subjective tests. For *noise-separations* in the range −30 dB to −50 dB, the distortion signal resembled randomly fluctuating impulses. This type of signal made the PPM pointer indication fluctuate by up to 20 dB and caused subjective estimates of the pointer indication to be somewhat unreliable. More reliable readings were possible with the r.m.s. meter, as it had an integration time of about 10 seconds, and gave a steady indication when monitoring the distortion signal.

With a test-noise power of 3 dBm, an r.m.s. *noise-separation* of approximately −35 dB was measured for each line receiving amplifier, but the *noise-separation* measured for the saturating transformer was very much greater (about −60 dB).

Ideally, a distortion test should give the same objective measurement for all non-linear circuits which produce equal subjective impairment. In order for the *noise-separation* test to meet this requirement more closely, the bandwidth of the test signal was adjusted so that the saturating transformer had a *noise-separation* of −35 dB with a test power of 3 dBm. This was achieved by reducing the upper

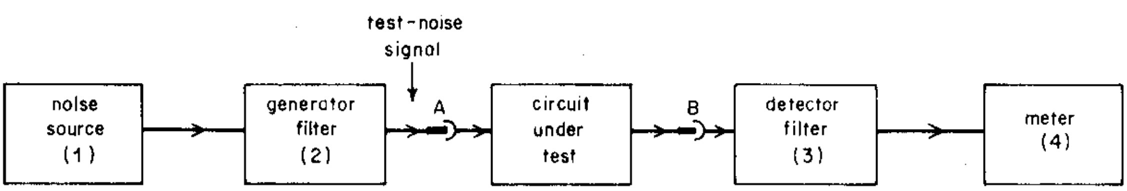
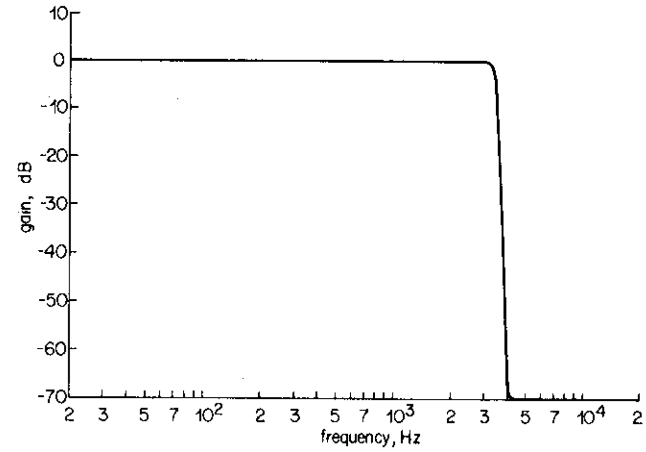
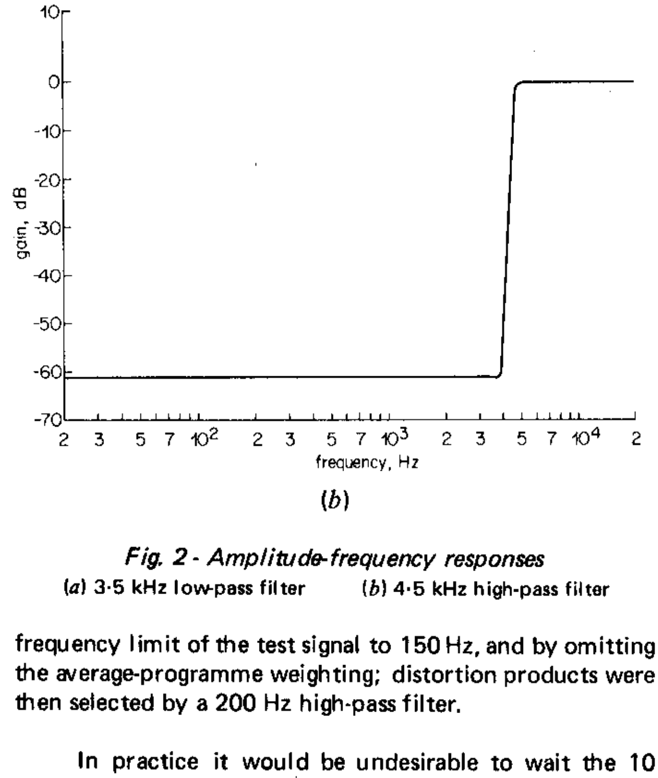


Fig. 1 - Noise-separation measurement



(*a*)



(*b*)

Fig. 2 - Amplitude-frequency responses
(*a*) 3·5 kHz low-pass filter (*b*) 4·5 kHz high-pass filter

frequency limit of the test signal to 150 Hz, and by omitting the average-programme weighting; distortion products were then selected by a 200 Hz high-pass filter.

In practice it would be undesirable to wait the 10 seconds or more required with the r.m.s. meter for each *noise-separation* reading, so the statistical nature of the test-noise signal was altered in order to make reliable measurements possible with a peak or quasi-peak reading meter. The random noise source was replaced with a pseudo-random noise source. This pseudo-random noise source was chosen to have a sequence length of 65,535 and a spectrum with components spaced regularly in frequency by 4·6 Hz.

The preliminary investigations have indicated that in order to take sufficient account of non-linear distortion which predominates at low frequencies, the *noise-separation* method requires either (a) a test-noise signal having two or more bands of noise, or (b) two single-band test-noise signals.

For comparison with the *noise-separation* measurements described above the total harmonic distortion of a sine-wave having a power of 10 dBm at point A, was also measured for each circuit. For programme signal levels corresponding to just perceptible impairment, t.h.d. measured at 1 kHz for the two line receiving amplifiers was −47 dB and −32 dB, and for the saturating transformer was −15 dB at 100 Hz and −64 dB at 1 kHz.

On the basis of the informal subjective data obtained in the preliminary tests, the *noise-separation* method for measuring distortion gives a better estimate of the subjective threshold of impairment due to non-linear distortion than provided by the total harmonic distortion method. A formal investigation was therefore conducted, as described in Section 3, to provide data relating subjective impairment to programme signal level for more critical programme items and a wider range of non-linear circuits than used in the preliminary investigations.

3. An assessment of programme quality impairment due to four non-linear circuits: formal investigations

3.1. Scope of subjective tests

The purpose of the subjective tests was to assess the programme-quality impairment which was experienced when circuits were over-driven, and to relate this to the level of the programme signal. No amplitude compression was applied to the programme signals, and programme peaks registered no more than '6' on a PPM at a point in the circuit where the line-up tone read '4'.

It was intended that the tests should explore impairment levels in the range of grades 1 to 4 on the six point scale[7] shown in Table 1. Three levels of impairment were investigated for each combination of circuit and programme item.

TABLE 1

The 6-Point Subjective Impairment Scale

Grade	Impairment
1	Imperceptible
2	Just perceptible
3	Definitely perceptible but not disturbing
4	Somewhat objectionable
5	Definitely objectionable
6	Unusable

3.2. Choice of test programme material

Several factors influenced the final selection of programme material used for these tests. For convenience, and for pracitcal reasons, it was necessary to record all the test excerpts on magnetic tape. The noise and distortion levels of these recordings had to be low enough to avoid masking any distortion introduced by the circuit under test. To ensure this, it was essential that the programme signals did not undergo the record-replay process more than twice. Further, a companding system was used with the

tape-recorder to give a better signal-to-noise ratio than normally obtainable.

Several examples of critical programme material were employed to find if an objective measurement of distortion could give equal estimates of subjective impairment for each of the programme items chosen. The peak-programme signal level of each of these programme items was measured (a) using a PPM, and (b) using a storage oscilloscope. This yielded the quasi-peak and true-peak values for the programme signals. In Table 2, these results are given relative to the peak value of the line-up tone (dBO), and they are discussed later in Section 5.6.

TABLE 2

Peak Programme Levels

Programme Excerpt	Peak Signal Level Measured By: (a) a PPM	(b) a Storage Oscilloscope
Male Voice Choir	8 dBO	12·6 dBO
Male Speech	8 dBO	13·2 dBO
Five Note Piano Scale	8 dBO	10·6 dBO

Various piano, speech and choral programme excerpts were critical of programme signal distortion at peak levels. These items were considered to be typical of broadcast material and were used for the subjective tests. The duration of each excerpt had to be long enough for listeners to give reliable assessments, but the total duration of the final presentation tape had to be short enough to avoid listener fatigue. Each programme excerpt was about 20 seconds long and separated by about 5 seconds of silence.

3.3. Conduct of subjective tests

Several points were included in the planning of the subjective tests to enable the reliability and repeatability of listeners' assessments to be determined. Each test presentation comprised a nominally unimpaired excerpt followed by the impaired version. This format was used to assist the listener in rejecting any programme distortion which had not resulted from the circuit under test. In order to restrict the duration of each test session to about 20 minutes, the investigation was divided into two groups, as shown in Table 3.

For each session there were 24 comparisons, which included a selection of test conditions that were presented twice to check the consistency of listeners' assessments. These test conditions were presented in a random order. Nevertheless, it was recognised that at the beginning of each test session the judgement of some listeners might have been less critical than at half-way through the session. In addition, they could have been unduly critical at the end of the session, as they would then have become more familiar with the nature of the impairment. Hence two test sessions were required in each group, the second session being used to present the test conditions of the first in a different running order. This new running order can be explained by way of the following example. If the first test session used test conditions numbered from 1 to 20 then the order in which these would be presented in the second test session would be numbers 11 – 20 followed by numbers 1 – 10. The results of these two test sessions were compared and those with the lower standard deviation were selected for future use.

TABLE 3

The Subjective Tests

PROGRAMME EXCERPTS (A, B, C.)

A : Male Voice Choir
B : Male Speech
C : Five Note Piano Scale

CIRCUITS TESTED (1, 2, 3, 4.)

1 : Integrated circuit operational amplifier type '741'
2 : AM7/4 BBC Line receiving amplifier
3 : GPA/4A BBC Line receiving amplifier
4 : GPA/4A BBC Line receiving amplifier with low cathode emission

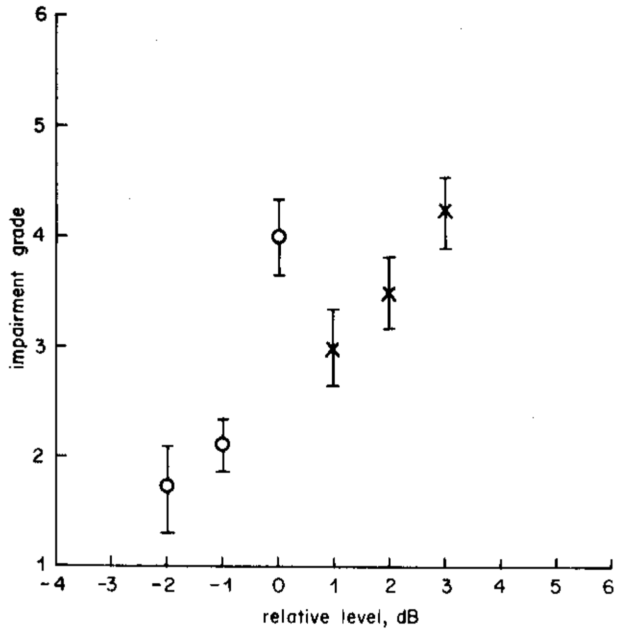
ARRANGEMENT

First Group	A1,	A2,	B1,	B2,	C1,	C2.
Second Group	A3,	A4,	B3,	B4,	C3,	C4.

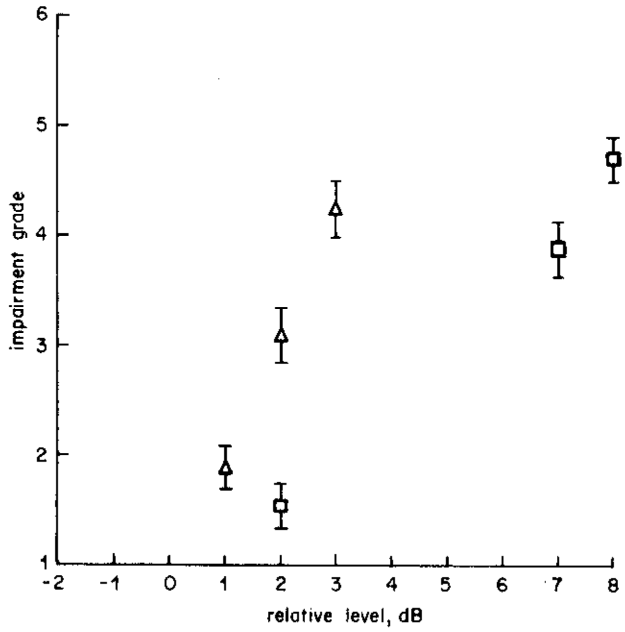
For each subjective test, listeners were asked to assess impairment in terms of the 6-point impairment scale given in Table 1. They were asked to assess the absolute impairment of the second presentation in each case bearing in mind that the first presentation should be regarded as being unimpaired (grade 1). Each test condition was separated by a five second pause during which the listeners were required to record their assessments. Details of the presentation order of the test conditions for each test session are given in the Appendix.

The subjective tests were carried out in a listening room which had acoustics similar to those of an average domestic living room. Each test condition was judged by eleven listeners, each experienced in assessing sound quality, but the tests were arranged so that a maximum of six listeners attended each session. The programme excerpts were reproduced at a sound level of approximately 85 dBA on a BBC monitoring loudspeaker type LS5/5.

The non-linear circuits used for these subjective tests are described in Table 3. It was essential that the operating point used for each circuit could be defined accurately. A reference condition was chosen such that a 1 kHz signal level of +10 dBm applied at point A in Fig. 1 produced a t.h.d. of −37 dB; the input gain in each circuit being adjusted as necessary to obtain precisely this amount of distortion (the output gain being adjusted to give zero gain overall under linear operating conditions). The term 'relative level' is used to indicate the input gain of each non-linear circuit relative to the reference condition.



(a)



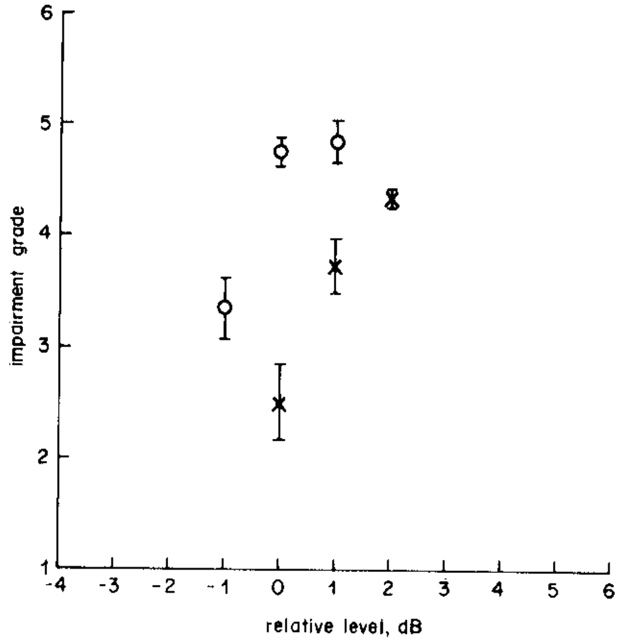
(b)

Fig. 3 - Average impairment grade versus relative level for 'male-voice choir' test

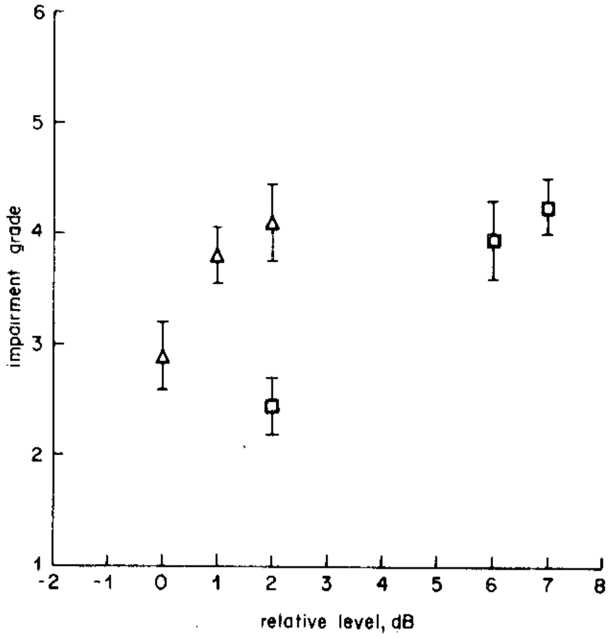
(a) with circuits 1 and 2   *(b)* with circuits 3 and 4

O circuit 1   X circuit 2   △circuit 3   □circuit 4

I twice standard error



(a)



(b)

Fig. 4 - Average impairment grade versus relative level for 'male speech' test

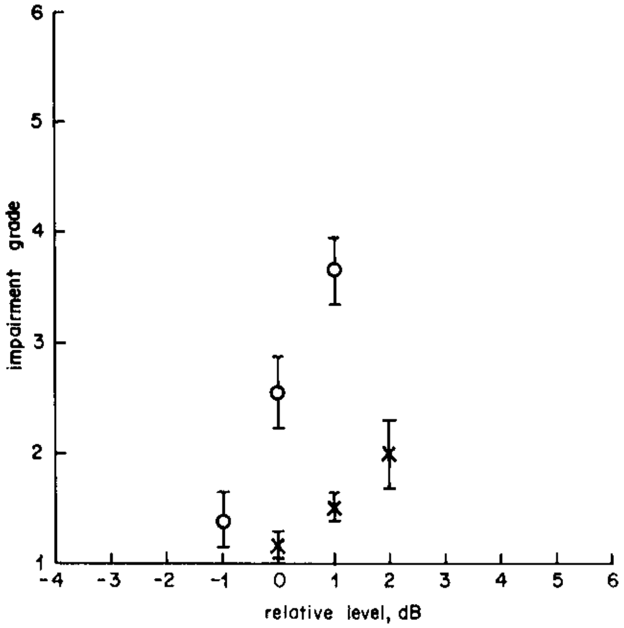
(a) with circuits 1 and 2   *(b)* with circuits 3 and 4

O circuit 1   X circuit 2   △circuit 3   □circuit 4

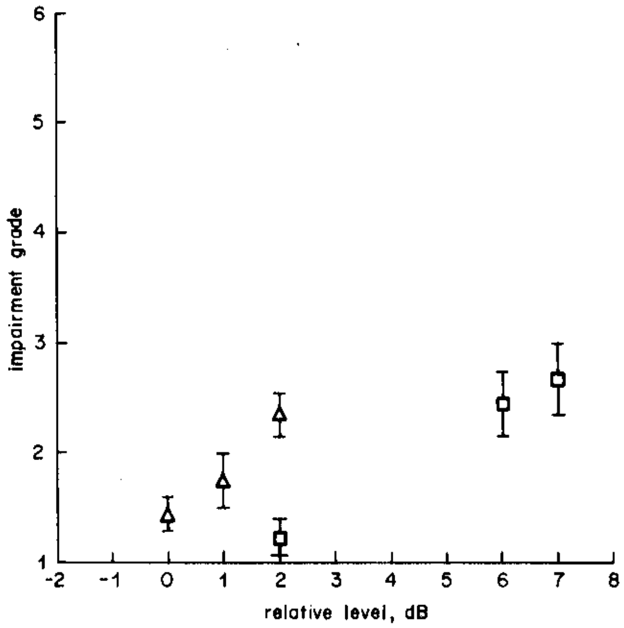
I twice standard error

3.4. Results of the tests

The results of these tests were analysed statistically as stated in Section 3.3 and are presented in Figs. 3, 4 and 5 where mean impairment grade is plotted against relative level. In this report, the least 'tolerant' circuit is the one for which the rate of change of impairment with relative level is greatest. Programme items are classified as follows: the most 'critical' programme item will denote that programme which requires the least relative level to give grade 2 impairment.



(*a*)



(*b*)

Fig. 5 - Average impairment grade versus relative level for 'five-note piano scale' test

(*a*) with circuits 1 and 2 (*b*) with circuits 3 and 4
O circuit 1 X circuit 2 △circuit 3 □circuit 4
I twice standard error

3.4.1. Male-voice choir excerpt

Figs. 3(*a*) and 3(*b*) show the results obtained with programme 'A', male-voice choir. Circuits 1, 2, 3 and 4 of Table 3 have corresponding rates of change of impairment with relative level of: 2 grades/dB, ½ grade/dB, 1 grade/dB and ½ grade/dB, respectively.

3.4.2. Male speech excerpt

The results for programme 'B', male speech excerpt, are given in Figs. 4(*a*) and 4(*b*), and in the same context as used under Section 3.4.1 the rates of change are 1½ grades/dB, 1¼ grades/dB, 1 grade/dB and ¼ grade/dB respectively. The relative levels at which impairment began were generally lower than for test programme 'A'. so programme 'B' is more critical.

3.4.3. Five-note piano scale

In Figs. 5(*a*) and 5(*b*) which refer to programme 'C', a five-note piano scale, the rates of change are 1¼ grades/dB, ½ grade/dB, ½ grade/dB and ¼ grade/dB in the same order as used previously. The relative level at which impairment started was higher than for programmes 'A' or 'B'. Hence programme 'C' is least critical of distortion at peak signal levels.

3.4.4. Summary of results

From these tests the most critical programme excerpt was male speech and the least critical was the five-note piano scale. In addition, from a comparison of the rates of change of impairment with 'relative level', circuit 1 is the least 'tolerant' and circuit 4 the most 'tolerant'. Unimpaired programme was judged by listeners to be between grades 1 and 1·2. Hence listeners had on the whole been able to reject any residual distortion present in the programme item, and to base their assessment on the distortion deliberately introduced by each non-linear circuit. The standard error for mean estimations of subjective impairment varied from ±0·1 to ±0·35 of a grade: about 68% of mean estimations by other experienced listeners would be expected to lie within these limits.

The results given in Figs. 3, 4 and 5 are used in later sections as a basis for judging the relative success of various objective distortion measuring techniques in indicating the subjective impairment to be expected for various programme signal levels.

4. Further investigation of test-noise signals

4.1. Development of a single-band test-noise signal

The 3·5 kHz bandwidth test-noise signal described in Section 2.3 was re-assessed using the results of the subjective tests described in Section 3. These tests included a wider range of non-linear circuits and more critical programme material than was used initially, but in order to limit the time of the investigation, no subjective data was sought for distortion caused by a saturating transformer. This new data showed that the correlation between subjective assessment and objective measurement of distortion was not as good as had been suggested in the preliminary tests, which used the same single-band test-noise signal. As a result further objective experiments were conducted to develop a *noise-separation* test which would be more suitable for estimating the impairment caused by over driven amplifiers.

The test-noise signal used in the preliminary tests was used as a starting point. This signal was formed by filtering pseudo-random white noise to limit its bandwidth to 3·5 kHz, as shown in Fig. 2(*a*). No average-programme spectrum weighting network was employed for the test-noise signal. Instead, a variety of spectral weightings were applied to the test signal and to the distortion products, using two variable equalisers. The general block diagram shown in Fig. 1 was used as a basis for the development of the test-noise signal.

Noise-separation tests were conducted for each type of detector weighting and best correlation with subjective impairment was obtained with no generator weighting. Four meter circuits were used for these tests, (a) a PPM, (b) a modified Niese meter (previously known as a modified OIRT meter[8]) (c) a PPM used with CCIR 1970 noise weighting,[9] and (d) a modified Niese meter used with CCIR 1970 noise weighting.[9]

An impairment of the most critical programme item of grade 2 on the six-point scale was used to compare the results of each combination; ideally each non-linear circuit should give equal *noise-separations* for the same impairment. The optimum combination of meter circuit and detector weighting was meter (b) with a 12 dB per octave high frequency 'emphasis' network; good results were obtained using meter (a) with no detector weighting. *Noise-separations* measured using meters (c) and (d) did not give such good correlation with subjective impairment. Between the four circuits the range of *noise-separation* for grade 2 impairment was 12 dB using the best form of single-band *noise-separation* test, and 16 dB using the next best one. These figures compare with a range of 26 dB in the total harmonic distortion (t.h.d.) measured at 10 dBm and 1 kHz.

This phase of the work was successful in providing better correlation with subjective impairment from a *noise-separation* test than from a t.h.d. test. The results encouraged further experiments with test-noise signals, and several variants, using two or more bands of noise were therefore studied.

4.2. Multi-band test-noise signals

Variations of non-linearity with frequency could be more accurately measured by using a test signal comprising a comb of frequency bands. To instrument this, two complementary comb filters would be constructed, one to form the test signal, and the other to select distortion products in the spectral gaps of the test signal.

A simplified version of this idea has been reported by the NHK.[4] They used a test signal with a power spectrum similar to that of an average programme, removed from it a band of noise ¾ octave wide, and selected distortion products in this spectral gap with a ¼ octave wide filter of the same centre-frequency. The centre-frequencies of the two filters were altered in ½ octave steps to cover most of the audio frequency range, and a plot of 'noise-separation' against frequency was drawn. The NHK examined non-linearity in a variety of sound broadcasting circuits using this technique, but did not propose any relationship between the objective measurement and the corresponding subjective impairment.

Instrumentation of the complementary comb filter system was considered to be beyond the scope of the present investigation, and a non-linearity test which involved many measurements was not considered to be the most appropriate for routine test purposes.

Accordingly, test-noise signals containing two, three or four bands of noise were examined. Two 1/3 octave filter sets were available for forming the band-elimination filters of the test-signal generator, and the band-pass filters of the detector. Unfortunately none of these arrangements were successful as the out-of-band attenuation provided by the 1/3 octave filter set proved to be insufficient.

The next step in the investigation was to consider a simpler form of test signal which used only two bands of noise.

4.3. General form of test signal used in final tests

Several filters were designed in order to provide two

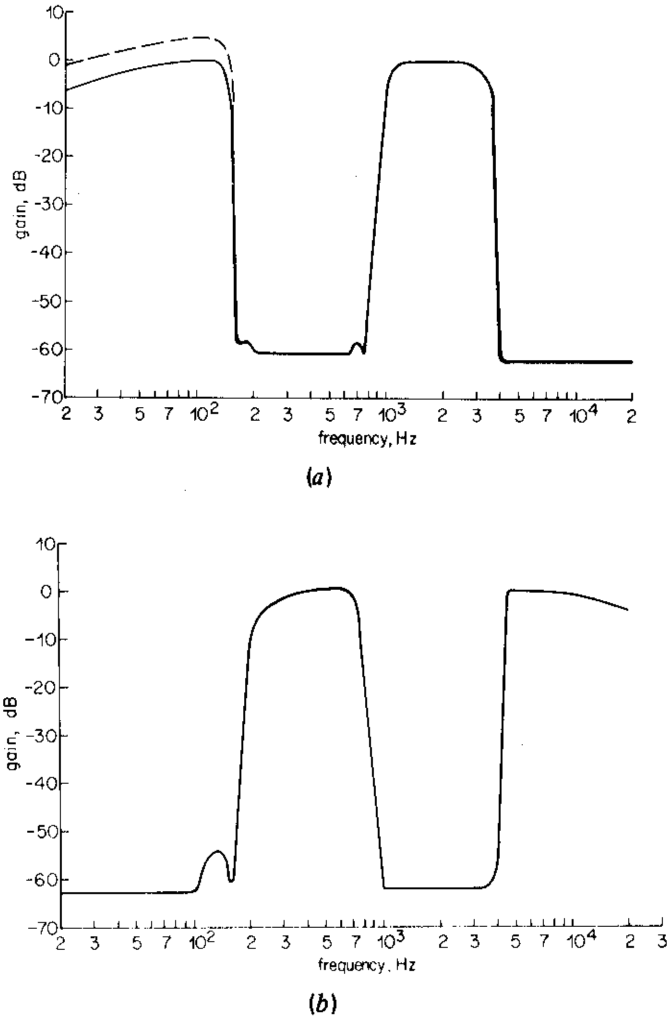


Fig. 6 - Amplitude-frequency responses
(*a*) the generator filter (*b*) the detector filter

bands of noise for the test signal: 20 to 150 Hz and 1 kHz to 3·5 kHz approximately. The amplitude-frequency response of the filter arrangement is shown by the full line in Fig. 6(*a*). Distortion products in the range 200 to 750 Hz and 4·5 kHz to 20 kHz were then measured using the filter response shown in Fig. 6(*b*).

Various combinations of meter circuit and spectral weighting were tried with this test-noise in the same way as described for the single-band test-noise in Section 3.2. Promising results were obtained, and a series of experiments were conducted to find optimum values of test-signal power, test-signal spectral weighting and distortion-product spectral weighting, and also the best type of meter to use.

5. Means of improving the correlation between subjective impairment and measurement of noise-separation

5.1. General

A series of experiments was arranged to optimise the general form of *noise-separation* test described in Section 4.3. Four experiments were carried out in the order presented by the following sub-sections. This order was chosen so that each experiment could best employ the results of those preceding it. The four meters described in Section 4.1 were used in these experiments, but for clarity no reference is made to this unless better correlation was obtained with a particular type of meter.

5.2. Choice of spectral weighting and relative power level for the test signal

The correct choice of power level for the test signal is of primary importance for a *noise-separation* form of non-linearity test. For reference purposes, this power level is quoted relative to the programme line-up level, as described in Section 2.2. Several power levels were used with each type of weighting employed with the test signal, in order to find the combination which gave best results for the most critical programme item. Weighting was applied to the test signal in two ways: (a) by varying the level of one band of noise relative to the other, and (b) by applying various frequency-emphasis characteristics to each band.

In practice the best correlation was obtained without applying frequency-emphasis to the noise bands, but with the low-frequency band of noise 5 dB higher in level than the high-frequency band as indicated by the dashed line in Fig. 6(*a*). The optimum level of this composite test-noise signal, at point A in Fig. 1, was found to be 3 dBm.

5.3. Spectral weighting of distortion products

The effect of applying spectral weighting to the output of the detector filter was investigated. The types listed in (a) and (b) of Section 5.2 were tried, but best results were obtained with no weighting of the distortion products.

5.4. Addition of random noise to the test signal

The effect was studied of adding random noise to the test signal. When this random noise was added it was found to improve estimates of impairment in the range of grades 3 to 1 by reducing the spread in the value of *noise-separation* which corresponded to a given impairment grade. The optimum improvement was obtained when the added noise, measured at point A in Fig. 1, was −43 dBm.

5.5. Choice of meter for measuring noise-separation

Noise-separation measurements were made using the four meters described in Section 4.1. This revealed that the type of meter influenced the reliability of the test for estimating low grades of impairment, and hence the performance of each meter was judged by measuring the *noise-separation* for which each non-linear circuit produced an estimated grade 1·5 impairment of the most critical programme item. The merit of each meter was related to the spread of these results. The PPM gave the best results: only a 2 dB variation for the four circuits. The other meters gave spreads of up to 5 dB at an estimated grade 1·5 impairment. The level of −43 dBm chosen for the added noise was the best compromise for all meters, but it would be possible to reduce the spread for the Niese meter by using a lower level of added noise.

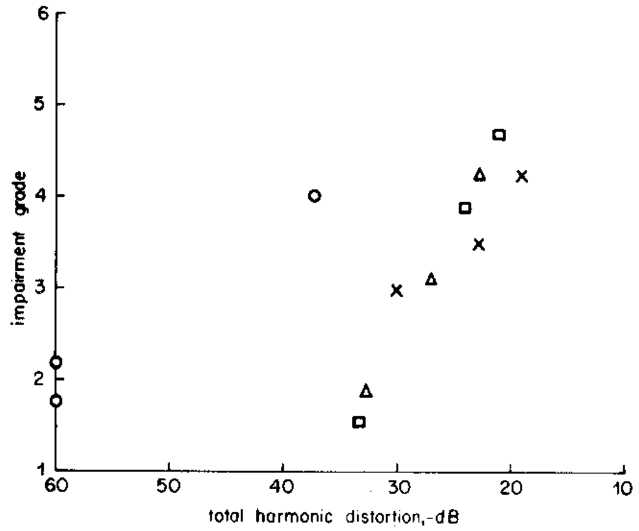
At low impairment grades, and for the most critical test programme, only the highest amplitude peaks of the test signal may explore the non-linearity. In this situation the distortion signal has a large peak-to-mean ratio. The PPM gives more weight to this type of signal than the Niese meter, and this may be the reason for it giving better results.

5.6. Programme level meters: their influence on the accuracy of the noise-separation test

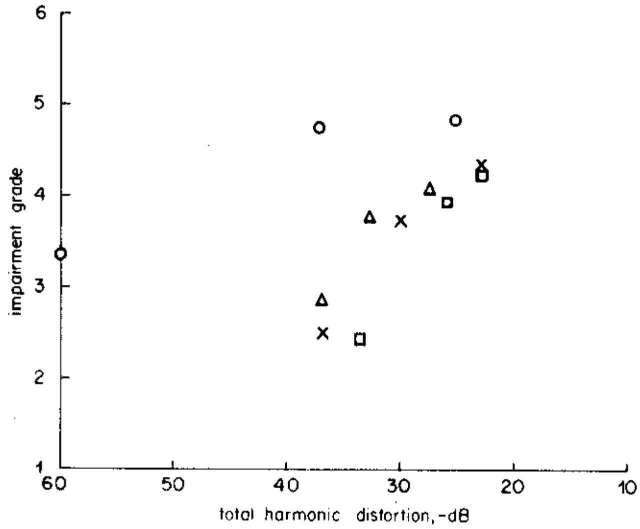
The *noise-separation* test described in the preceding sub-sections has been optimised for the most critical programme item, though it gives useful estimates of impairment for less critical programme items. This optimum form was developed for use in sound signal distribution circuits where programme signal level is monitored by a standard PPM.[6]

The effect of using a programme meter that gives more weight to the true peak signal level was investigated briefly. Peak programme levels were measured (a) with a standard PPM, and (b) with a storage oscilloscope, and the results recorded in Table 2. Previously, *noise-separation* has been measured with a test power of 3 dBm at point A in Fig. 1, where the programme line-up level was 0 dBm, and nominal peak programme level 8 dBm. For this brief investigation, *noise-separation* was measured with a power level of x dBm less than the true peak programme levels given in Table 2. Several values were tried for x to see if there was an optimum.

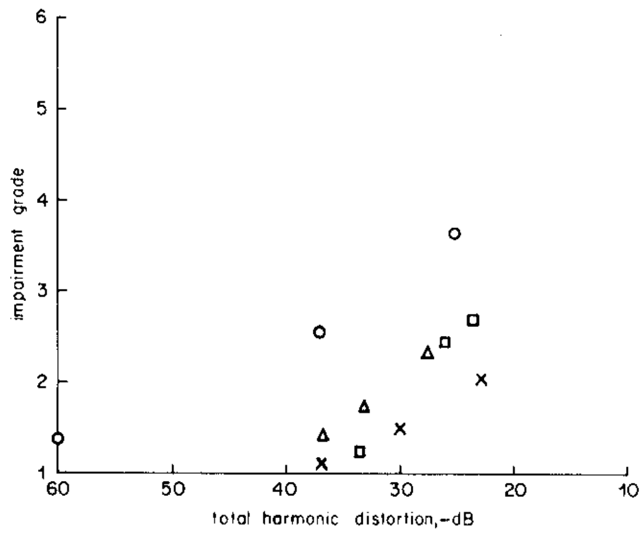
These tests indicated that impairment estimates with the *noise-separation* test would be less dependent on pro-



(*a*)



(*b*)



(*c*)

Fig. 7 - Total harmonic distortion versus mean subjective impairment grade

(*a*) 'Male-voice choir' test (*b*) 'Male speech' test

(*c*) 'Five-note piano scale' test

O circuit 1 X circuit 2 △circuit 3 □circuit 4

gramme material if the test-noise power level could be defined relative to a special quasi-peak estimate of the programme signal level. This estimate would be somewhere between the true-peak value, and the quasi-peak value indicated by a standard PPM: this could possibly be obtained with a modified PPM.

If programme levels are monitored with this optimum type of PPM rather than a standard PPM, it is likely that impairment estimates made by a noise-separation test would be as accurate, but less programme dependent.

6. A comparison of the reliability of two objective techniques for estimating subjective impairment

The introduction to this report has mentioned several techniques for measuring distortion. Two of these are now considered further: (a) the total harmonic distortion test, and (b) the optimised form of *noise-separation* test. The subjective data given in Section 3 were used to assess the merit of each distortion test. The following sub-sections contain graphs of subjective impairment versus objective measurement, and these form a basis for the comparison of each technique.

6.1. Total harmonic distortion

Figs. 7(*a*), (*b*) and (*c*) relate to programmes A, B and C respectively of Table 3. Each figure contains four plots of impairment versus t.h.d.; one plot for each non-linear circuit. T.H.D. is measured at 1 kHz and 10 dBm, referred to point A in Fig. 1.

Fig. 7(*b*) shows the results for the male speech excerpt, the most critical programme item. The four plots in Fig. 7(*b*) show that the t.h.d. test is inadequate for estimating the impairment caused by the four circuits: partly because of the divergence of the plots with decreasing impairment, but mainly because of the large difference between the t.h.d. measured for circuit 1 and the three other circuits at equal subjective impairment levels.

The plots of Figs. 7(*a*) and (*c*) follow the same general trend as 7(*b*), though the divergence of these plots is less. There is, however, a consistent failure to give sufficient weight to the impairment caused by circuit 1. This is shown by the overall spread in measurements of t.h.d. for grade 2 impairment: between 20 and 28 dB for graphs 7(*a*), (*b*) and (*c*).

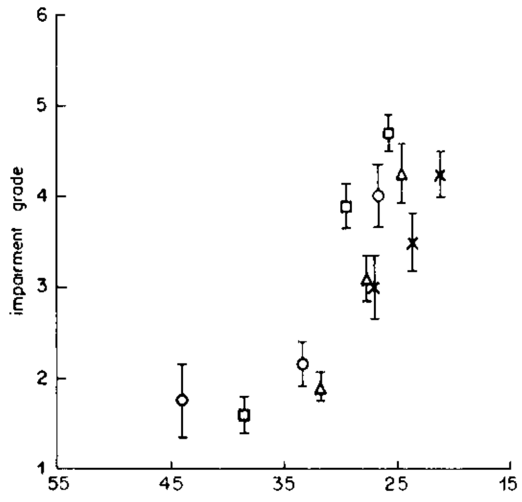
The gradients of the plots in Fig. 7 vary over a wide range of values, and vary most in Fig. 7(*b*); from 6 to 26 dB of t.h.d. per subjective impairment grade. Hence it is not possible to make an estimate of the rate of change of impairment with t.h.d. that would hold for all four circuits.

6.2. Noise-separation

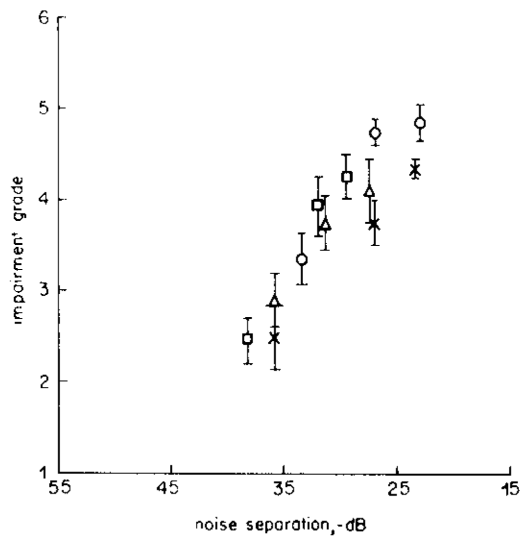
Sections 4 and 5 described various types of *noise-separation* test. The most accurate of these requires (a) a test signal with two bands of noise, and (b) a PPM to measure the distortion products. The optimised form of the *noise-separation* test is used for the plots contained in Figs. 8(*a*), (*b*) and (*c*), for three programme items respectively.

If the four plots in Fig. 8(*b*) are extrapolated, they converge near grade 1 impairment and 46 dB *noise-separation* and have very similar gradients. There is good agreement between the *noise-separation* and the subjective impairment: the four plots differ from each other by only 2 dB at grade 2, and 6 dB at grade 4. The gradients of the four plots in Fig. 8(*b*) vary from 5 to 7 dB of *noise-separation* per grade of subjective impairment, and, on average, the rate of change is 6 dB per grade.

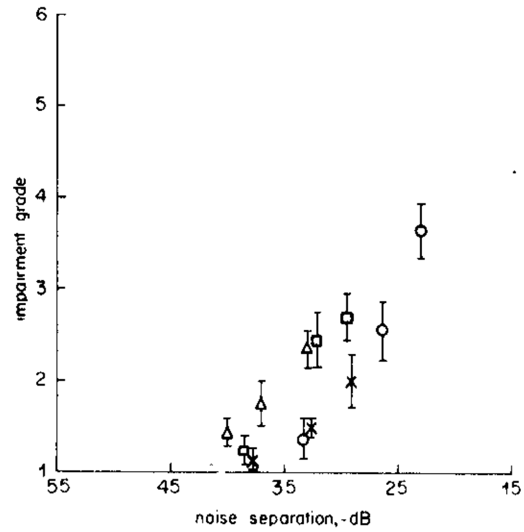
The plots in Figs. 8(*a*) and (*c*) are not as ordered as those in Fig. 8(*b*), but this fact is of minor significance as (*a*) and (*c*) relate to the less critical programme excerpts. The plots in Figs. 8(*a*) and (*c*) might have become more ordered if the programme signal levels had been adjusted using the special type of PPM discussed in Section 5.6. It is likely that the accuracy of impairment estimates made from the *noise-separation* test would then have been improved for the less critical programme items.



(*a*)



(*b*)



(*c*)

Fig. 8 - Noise-separation versus mean subjective impairment grade

(*a*) 'Male-voice choir' test (*b*) 'Male speech' test

(*c*) 'Five-note piano scale' test

O circuit 1 X circuit 2 △ circuit 3 □ circuit 4

I twice standard error

6.3. The comparison

The total harmonic distortion technique seems useful for rough estimates of impairment caused by circuits in which distortion does not increase rapidly with signal level. It might be suitable for routine tests where only those types of circuit were employed, and where an uncertainty of at least ±0·8 grade for each estimate is acceptable.

The *noise-separation* test was successful in estimating the degrees of impairment caused by all four circuits. This test is more useful than t.h.d. as thresholds of impairment can be estimated accurately. Impairment of the most critical test programme excerpt could be estimated with an accuracy of better than ±0·2 grade for impairments less than grade 2½ and within ±0·5 grade for grade 4.

7. A proposal for a routine test for non-linear distortion

7.1. Scope of the test

A *noise-separation* test is proposed for estimating programme quality impairment due to sound signal distribution circuits which produce distortion at crest values of the programme signal. Distortion of this nature can arise if the circuit becomes faulty, or if the programme signal level has been incorrectly adjusted.

7.2. The test

Noise-separation is measured using the method described in Section 2.2. The test-noise signal comprises two frequency-bands of pseudo-random noise; distortion products are selected by a special filter and their level measured with a standard PPM. Further details of this test-noise signal and special filter are given in Section 5.

Routine tests for non-linear distortion should employ a test-noise signal power of 3 dBm measured at a point of zero relative level. A standard PPM calibrated to read PPM4 on programme line-up tone would on this test-noise signal read PPM6.

The waveform of the test-noise signal had unequal maximum values of positive and negative excursion about the mean value of the signal; this was a consequence of the particular pseudo-random sequence employed. It was therefore necessary to make two measurements of *noise-separation* for each estimate of impairment: one with and one without a phase reversal of the test-noise signal. In practice, the *noise-separation* value to be used is the one with the smaller modulus. In general, the two values are similar; the esception arises when the non-linear circuit produces 'asymmetric clipping', and in particular where the amount of distortion increases markedly with signal level.

To facilitate *noise-separation* tests in service, the problem of phase reversing the test-noise signal could be easily resolved by arranging for the test-noise generator to provide a periodic phase-reversal of the test-noise signal, possibly at the end of each pseudo-random sequence.

7.3. Accuracy of the results

The value of routine non-linearity tests on sound programme circuits is related to the accuracy with which the measurement indicates the impairment to be expected on representative programme material transmitted by the circuit under test. Considering tests carried out on male speech, the significance of a *noise-separation* measurement may be estimated as indicated below.

TABLE 4

Comparison of proposed method with harmonic distortion method for estimating subjective grade (interpolated results for male speech programme)

Subjective grade (Table 1)	Method	Test Circuit			
		1	2	3	4
2·5	Noise-separation (−dB)	39*	37	39*	39
	T.H.D. (−dB)	60*	37	40*	34
3·5	Noise-separation (−dB)	34	30	34	34
	T.H.D. (−dB)	58	31	34	32
4·5	Noise-separation (−dB)	29	30	24*	28
	T.H.D. (−dB)	42	22	22*	21*

* extrapolated result

8. Conclusions

Listening tests have been conducted to assess the programme quality impairment produced by four non-linear circuits, and several objective methods of measuring their non-linearity have been investigated. The success of each method of measurement was assessed from the degree of correlation between the measured distortion and subjective assessments of impairment of programme.

A *noise-separation* method is proposed by which routine tests can be made for non-linearity in various types of circuit. This method gives much more consistent results relating the measured distortion with its subjective effect than total harmonic distortion. The variation in total harmonic distortion which related to a given subjective impairment was in fact so great as to make this method totally unsuited for estimating subjective impairment, unless the type of circuit non-linearity is also known. This can be seen from Figs. 7 and 8 and Table 4.

The *noise-separation* is measured at a test-noise level whose value is of necessity defined relative to some 'measure' of the programme signal level. For convenience this investigation used the programme line-up level as a 'measure' of the programme signal level. A brief series of experiments was conducted to find out if the use of this 'measure' was consistent with obtaining the best performance from the *noise-separation* test.

These experiments indicated that the performance of

the *noise-separation* test could be improved by using a different 'measure'; estimates of impairment could be made less dependent on the programme material. This new 'measure' was a quasi-peak estimate of the programme signal level, and was somewhere between the level registered by a standard PPM, and the true peak signal level. It is possible that this optimum quasi-peak value might be indicated by a modified PPM having a faster attack time.

9. Recommendations

The main subjective tests described in this report were limited to assessing programme impairment in circuits where the non-linearity was not a function of frequency. Further subjective tests should be conducted to include situations where non-linearity occurs predominantly at high or low audio frequencies. Some examples can be provided by the following circuits: (a) a saturating transformer giving distortion which increases at low frequencies, and (b) an over-modulated f.m. transmission link where high frequencies are more vulnerable to distortion owing to the f.m. pre-emphasis characteristic.

The *noise-separation* test proposed in Section 7 should be suitable for assessing low frequency distortion, but would have to be modified before it could detect a non-linearity which was confined to the higher audio frequencies. One solution would be to include a further band of noise somewhere in the range of 5–15 kHz, about 1 kHz wide. Intermodulation components would then be detected in the lower pass-band of the measuring filter. This seems feasible as early work by Dånes[5] used a 1/3 octave band of noise centred on 8 kHz in the measurement of intermodulation distortion in an f.m. transmission link.

A test-noise signal explores circuit non-linearities over a wide dynamic range. This should make it useful for measuring the type of distortion which occurs in pulse-code modulated (p.c.m.) sound-signal distribution circuits, and in particular those p.c.m. networks which employ a companding system.

10. References

1. SHORTER, D.E.L. 1950. The influence of high-order products on non-linear distortion. *Electron. Engng.*, 1950, **22**, 4, pp. 152 – 3.

2. WIGAN, E.R. 1961. New distortion criterion. *Electron. Technol.*, 1961, **38**, 4 & 5, pp. 128 – 137, 163 – 174.

3. CCIR New Delhi, 1970, **V**, 2, Report 497.

4. NIKAIDO, S. *and* NITATORI, T. 1971. Non-linear distortion measurement of audio systems using noise signal. Proc. 7th International Congress on Acoustics, Budapest, 1971, **1**, 1, 19E 10, pp. 301 – 304.

5. DÅNES, A. 1963. A sibilant limiter for f.m. broadcasting systems. *EBU Rev.*, No. 77, Pt. A., February 1963.

6. British Standard No. 4297 : 1968.

7. CCIR New Delhi, 1970. Report 405-1.

8. GEDDES, W.K.E. 1968. The assessment of noise in audio-frequency circuits. BBC Research Department Report No. EL-17, Serial No. 1968/8.

9. CCIR New Delhi, 1970. Recommendation 468.

11. Appendix

Presentation Order of Test Conditions

The test conditions of Table 3 were presented in a random order. Each condition presented an estimated quality impairment in the range of grades 1 to 4 on the six-point scale in Table 1, and related to one programme excerpt that had been processed by one circuit.

The notation used to describe each condition given below conforms to the following format:

letter (programme item)	*number* (circuit type)	at *number*	(impairment grade) (estimated by) (one listener)

11.1. Session 1

Group 1		Group 2	
Item	Condition	Item	Condition
1	C1 at 3½	1	C3 at 3½
2	C1 at 1	2	C3 at 1
3	C1 at 2½	3	C3 at 2½
4	C1 at 1½	4	C3 at 1½
5	B1 at 1	5	B3 at 1
6	B1 at 1½	6	B3 at 1½
7	B1 at 3½	7	B3 at 3½
8	B1 at 2½	8	B3 at 2½
9	A1 at 3½	9	A3 at 3½
10	A1 at 2½	10	A3 at 2½
11	A1 at 1	11	A3 at 1
12	A1 at 1½	12	A3 at 1½
13	B2 at 3½	13	B4 at 3½
14	B2 at 1	14	B4 at 1
15	B2 at 2½	15	B4 at 2½
16	B2 at 1½	16	B4 at 1½
17	C2 at 1	17	C4 at 1
18	C2 at 1½	18	C4 at 1½
19	C2 at 3½	19	C4 at 3½
20	C2 at 2½	20	C4 at 2½
21	A2 at 1	21	A4 at 1
22	A2 at 1½	22	A4 at 1½
23	A2 at 3½	23	A4 at 3½
24	A2 at 2½	24	A4 at 2½

11.2. Session 2

Session 2 presented the same test conditions as Session 1, but in a different order. The conditions in Session 1 were re-arranged so that those numbered from 1 to 24 in 11.1 were presented in the order 13 to 24 first, and 1 to 12 last.

SMW/AMM